

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2003, 23:44:32 ; Search time 1168.39 Seconds  
(without alignments)  
471.288 Million cell updates/sec

Title: US-09-913-524-32  
Perfect score: 34  
Sequence: 1 aggcctcggagggaacggcgtgcccatgccaaact 34

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	251	14	T29423 EST79588 Hu
2	34	100.0	364	10	AW951334 EST363404
3	34	100.0	501	9	AI128256 qc33h05.x
4	34	100.0	884	13	BI520936 603071335
5	34	100.0	899	9	AL568487 AL568487
6	34	100.0	924	9	AL540575 AL540575

c	7	32.4	95.3	599	9	AI220927
c	8	30	88.2	548	14	BQ013783
c	9	27.6	81.2	510	9	AI885743
	10	26	76.5	399	12	BE758121
	11	26	76.5	464	13	BI849730
	12	26	76.5	466	9	AA234854
	13	26	76.5	551	12	BE721322
	14	26	76.5	1074	13	BM460913
	15	26	76.5	1118	13	BM461343
c	16	25	73.5	641	13	BI186150
	17	24.4	71.8	640	12	BC872889
c	18	24.4	71.8	659	13	BI183158
c	19	24.4	71.8	705	13	BI182159
c	20	24.4	71.8	1172	13	BM461317
c	21	23.4	68.8	613	10	AW55078
c	22	23.4	68.8	649	13	BI181741
c	23	23.4	68.8	978	14	BM804780
c	24	23	67.6	629	13	BM665951
c	25	22.8	67.1	1129	13	BM461356
c	26	22	64.7	582	14	C87829
c	27	21.4	62.9	669	17	AG148285
c	28	21.4	62.9	1521	13	BM563318
c	29	21.2	62.4	466	10	AW475765
c	30	21.2	62.4	1009	17	CMS02E39
c	31	21	61.8	252	13	BM135954
c	32	21	61.8	384	9	AA219283
c	33	21	61.8	604	12	BG042907
c	34	20.8	61.2	227	9	AA030977
c	35	20.8	61.2	488	9	AA979198
c	36	20.8	61.2	538	10	BB769745
c	37	20.8	61.2	542	9	AI515111
c	38	20.8	61.2	597	9	AI512705
c	39	20.8	61.2	797	12	BG199330
c	40	20.6	60.6	740	13	BI1103224
c	41	20.6	60.6	843	13	BI332823
c	42	20.6	60.6	845	13	BI857862
c	43	20.6	60.6	920	14	BQ964399
c	44	20.4	60.0	278	9	AV088622
c	45	20.4	60.0	343	13	BG961308

ALIGNMENTS

RESULT 1  
T29423  
LOCUS EST79588 Human Placenta Homo sapiens cDNA similar to inhibin, alpha (HT:1847). mRNA sequence.  
DEFINITION T29423 251 bp mRNA linear EST 06-SEP-1995  
ACCESSION T29423  
VERSION T29423.1 GI:611521  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 251)  
AUTHORS Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A., Bult C.J., Lee N., Kirkness E.F., Weinstock K.G., Gocayne J.D., White O., Sutton G., Blake J.A., Brandon R.C., Chiu M.-W., Clayton R.A., Cline R.T., Cotton M.B., Earle-Hughes J., Fine L.D., Fitzgerald L.M., Fitzhugh W.M., Fritchman J.L., Geoghagen N.S.M., Glodek A., Gnehm C.L., Hanna M.C., Hedblom E., Hinkle Jr P.S., Kelley J.M., Klimek K.M., Kelley J.C., Liu L.-I., Marmaros S.M., Merrick J.M., Moreno-Palauques R.F., McDonald L.A., Nguyen D.T., Pellegrino S.M., Phillips C.A., Ryder S.E., Scott J.L., Saudek D.M., Shirley R., Small A.V., Spriggs T.A., Utterback T.R., Weidman J.F., Li F., Bednarek D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J., Dimke D., Feng P., Ferrie A., Fischer C., Hastings G.A., He W.-W., Hu J.-S., Greene J.M., Gruber J., Hudson P., Kim A., Kozak D.L., Kunsch C., Ji H., Li H., Weissner P.S., Olsen H., Raymond L., Wei Y.-F., Wing J., Xu C., Yu G.-L., Ruben S.M., Dillon P.J., Fannon M.R., Rosen C.A., Haseltine W.A., Fields C., Fraser C.M. and

Venter, J.C.  
 Initial Assessment of Human Gene Diversity and Expression Patterns  
 Based Upon 83 Million Basepairs of cDNA Sequence  
 Nature 377, 3-174 (1995)  
 96026280  
 Contact: Venter, JC  
 The Institute for Genomic Research  
 932 Clopper Rd, Gaithersburg, MD 20878  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: tdbinfo@tdb.tigr.org  
 For clone availability, additional sequence and expression  
 information related to this EST, please contact the TIGR Database  
 (tdbinfo@tdb.tigr.org)  
 Seq primer: M13 Reverse.

# FEATURES

source  
 1..251  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="ATCC (inhost):106213"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Human Placenta"  
 /note="Organ: placenta"

BASE COUNT 47 a 85 c 67 g 51 t 1 others  
 ORIGIN

Query Match 100.0%; Score 34; DB 14; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 0.037;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 94 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 127

RESULT 2  
 LOCUS AW951334 364 bp mRNA linear EST 01-JUN-2000  
 DEFINITION EST363404 MAGE resequencences, MAGB Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW951334  
 VERSION AW951334.1 GI:8141005  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 364)  
 AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt  
 J.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J., and  
 Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor  
 metastasis using a 19,200 element cDNA microarray  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: John Quackenbush  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 3528  
 Fax: 301 838 0208  
 Email: johnq@tigr.org  
 Plate: 26  
 Seq primer: Reverse.

# FEATURES

source  
 1..364  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MAGE resequencences, MAGB"  
 /note="Vector: pBluescriptSKm"

BASE COUNT 59 a 135 c 96 g 74 t  
 ORIGIN

Query Match 100.0%; Score 34; DB 10; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0.04;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 94 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 127

# RESULT 3

LOCUS A1128256 501 bp mRNA linear EST 27-OCT-1998  
 DEFINITION qc33h05.x1 Soares\_pregnant\_uterus\_NbHPU Homo sapiens cDNA clone  
 IMAGE:1711449 3' similar to gb:M13981 INHIBIN ALPHA CHAIN PRECURSOR  
 (HUMAN);, mRNA sequence.

ACCESSION A1128256  
 VERSION A1128256.1 GI:3596770  
 KEYWORDS EST.  
 SOURCE human.

# ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 501)  
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 832 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 433.

# FEATURES

source  
 1..501  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Soares\_pregnant\_uterus\_NbHPU"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/note="Organ: uterus; Vector: pT73-Pac; Site: 1; Not I;  
 Site: 2; Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer [5',  
 AACTGAGAGATTCGGCGCGCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization. Library  
 constructed by M. Fatima Ronaldo."

BASE COUNT 118 a 117 c 177 g 89 t  
 ORIGIN

Query Match 100.0%; Score 34; DB 9; Length 501;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 34  
 ||||||||||||||||||||||||||||||||||||  
 Db 445 AGGCCTCGGAGGACCGGCTGCCCATGCCCAACT 412

# RESULT 4

LOCUS B1520936 884 bp mRNA linear EST 29-AUG-2001  
 DEFINITION 603071335T1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5163273 3',  
 mRNA sequence.

ACCESSION B1520936  
 VERSION B1520936.1 GI:15345728  
 KEYWORDS EST.  
 SOURCE human.

# ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 884)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cdpbs-r@mail.nih.gov](mailto:cdpbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM11405 row: f column: 10  
High quality sequence start: 2  
High quality sequence stop: 726.

FEATURES  
source

1..884  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5163273"  
/clone\_lib="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."

BASE COUNT  
ORIGIN

175 a 241 c 327 g 151 t

Query Match  
Best Local Similarity  
Matches

100.0%; Score 34; DB 13; Length 884;  
100.0%; Pred. No. 0.048;  
34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 34  
|||||

DB

458 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 425

RESULT 5  
AL568487/C

LOCUS

AL568487 LTL\_FL002\_PL1 Homo sapiens cDNA clone CS0DE002YE07 3 prime  
899 bp mRNA linear EST 16-FEB-2001  
mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

Location/Qualifiers

1..899

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DE002YE07"

/clone\_lib="LTL\_FL002\_PL1"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime

cdna

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

34; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

1 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 34

|||||

DB

458 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 425

RESULT 6

AL540575

LOCUS

DEFINITION

AL540575 LTL\_FL002\_PL1 Homo sapiens cDNA clone CS0DE002YE07 5 prime

, mRNA sequence.

ACCESSION

AL540575

VERSION

AL540575.1 GI:12870853

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 924)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

Location/Qualifiers

1..924

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DE002YE07"

/clone\_lib="LTL\_FL002\_PL1"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime

cdna

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pcMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax : (1) 301 610 8371 Email :

[fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com>"

BASE COUNT

ORIGIN

184 a 229 c 318 g 159 t 9 others  
Query Match 100.0%; Score 34; DB 9; Length 899;  
Best Local Similarity 100.0%; Pred. No. 0.049;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 34  
|||||

DB 458 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 425  
|||||

RESULT 6  
AL540575

LOCUS

DEFINITION

AL540575 LTL\_FL002\_PL1 Homo sapiens cDNA clone CS0DE002YE07 5 prime

, mRNA sequence.

ACCESSION

AL540575

VERSION

AL540575.1 GI:12870853

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 924)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

Location/Qualifiers

1..924

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DE002YE07"

/clone\_lib="LTL\_FL002\_PL1"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime

cdna

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pcMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax : (1) 301 610 8371 Email :

[fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com>"

BASE COUNT

ORIGIN

153 a 307 c 299 g 165 t

Query Match 100.0%; Score 34; DB 9; Length 924;

Best Local Similarity 100.0%; Pred. No. 0.049;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 34  
|||||

DB 838 AGGCTCCGGAGAACCGCTGCCCATGCCCACT 871  
|||||

RESULT 7  
A1220927/C

LOCUS

DEFINITION

A1220927

599 bp mRNA linear EST 29-NOV-1998

qg08a11.x1 Soares\_placenta\_8to9weeks\_2NDHP8to9W Homo sapiens cDNA

clone IMAGE:1758908 3' similar to gb:M13981 INHIBIN ALPHA CHAIN  
 PRECURSOR (HUMAN);, mRNA sequence.  
 AI220927  
 AI220927.1 GI:3803130  
 EST.  
 HUMAN  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 599)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Insert Length: 761 Std Error: 0.00  
 Seq primer: -40up from Gibco  
 High quality sequence stop: 379.  
 Location/Qualifiers  
 1..599  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1758908"  
 /clone\_lib="Soares\_placenta\_8to9weeks\_2NBHP8to9W"  
 /dev\_stage="two placenta: one from 8 weeks and another  
 from 9 weeks post conception"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Ordan: placenta; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5',  
 TGTTACCAATCTGAAGGAGCGCGCGGATTTTTTTTTTTTTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Fatima Bonaldo."  
 BASE COUNT 135 a 142 c 203 g 115 t 4 others  
 ORIGIN  
 Query Match 95.3%; Score 32.4; DB 9; Length 599;  
 Best Local Similarity 97.1%; Pred. No. 0.16;  
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGGCCTCGGAGGACCGGCTGCCATGCCAACT 34  
 |||||||||||||||  
 DB 540 AGGCCTCGGAGGACCGGCTGCCATGCCAACT 507  
 |||||||||||||||  
 RESULT #  
 BQ013783/c 548 bp mRNA linear EST 26-MAR-2002  
 LOCUS  
 UI-1-BC1P-alb-g-08-0-UI-s1 NCI-CGAP.P13 Homo sapiens cDNA clone  
 UI-1-BC1P-alb-g-08-0-UI 3', mRNA sequence.  
 DEFINITION  
 BQ013783  
 ACCESSION  
 BQ013783.1 GI:19738684  
 VERSION  
 EST.  
 KEYWORDS  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 548)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Steven Brown  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.  
 Location/Qualifiers  
 1..548  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-1-BC1P-alb-g-08-0-UI"  
 /clone\_lib="NCI-CGAP.P13"  
 /tissue\_type="placenta"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Ordan: Placenta; Vector: pT73-Pac (Pharmacia) with  
 a modified polylinker; Site.1: EcoR I; Site.2: Not I;  
 NCI-CGAP.P13 is a subcloned cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are GA, AGGAA. For  
 additional information, contact: Bento Soares,  
[bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 TAG.LIB=UI-1-BC1P  
 TAG.TISSUE=placenta human 8 week  
 TAG\_SEQ=GA  
 BASE COUNT 122 a 123 c 185 g 118 t  
 ORIGIN  
 Query Match 88.2%; Score 30; DB 14; Length 548;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 CTCGGAGGAGCGGCTGCCATGCCAACT 34  
 |||||||||||||||  
 DB 548 CTCGGAGGAGCGGCTGCCATGCCAACT 519  
 |||||||||||||||  
 RESULT #  
 A1885743/c 510 bp mRNA linear EST 07-MAR-2000  
 LOCUS  
 w61a07.x1 NCI-CGAP.Bn25 Homo sapiens cDNA clone IMAGE:2429364 3',  
 similar to gb:M13981 INHIBIN ALPHA CHAIN PRECURSOR (HUMAN);, mRNA  
 sequence.  
 DEFINITION  
 A1885743  
 ACCESSION  
 A1885743.1 GI:5590907  
 VERSION  
 EST.  
 KEYWORDS  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 510)  
 NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute / National Institute of Neurological  
 Disorders and Stroke, Brain Tumor Genome Anatomy Project  
 (CGAP/BTGP), Tumor Gene Index  
 Unpublished (1998)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
 Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/btbp/image/image.html](http://www-bio.llnl.gov/btbp/image/image.html)

Trace considered overall poor quality  
 Insert Length: 1166 1Std Error: 0.00  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 1.

## FEATURES

Location/Qualifiers

1..510  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:2429364"  
 /clone\_lib="NCI CGAP Brn25"  
 /tissue\_type="neoplastic oligodendroglioma"  
 /lab\_host="DH10B"

/note=Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGAGGAGCGCCCATAGGTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 126 a 144 c 141 g 95 t 4 others

## ORIGIN

Query Match 81.2%; Score 27.6; DB 9; Length 510;  
 Best Local Similarity 88.2%; Pred. No. 7.7;

Matches 30; Conservative 0; Mismatches 4; Indels 0; Caps 0;

QY 1 AGGCTCGGAGGAGACCGGTCGCCATGCCCACT 34

||||||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 441 AGGCTCGGAGGAGACCGGTCGCCATGCCCACT 408

## RESULT 10

BE758121

LOCUS BE758121 399 bp mRNA linear EST 25-APR-2001

DEFINITION 212839 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE758121

VERSION BE758121.1

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 399)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCGACGACGACG

Plate: 66 row: I column: 16

Seq primer: ATTAGTGACACTATAG.

Location/Qualifiers

1..399

/organism="Bos taurus"

## FEATURES

Source

1..464

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SP08T6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from testis, thymus,

semitendonsus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

BASE COUNT 46 a 167 c 108 g 78 t

## ORIGIN

Query Match 76.5%; Score 26; DB 12; Length 399;

Best Local Similarity 85.3%; Pred. No. 27;

Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGGCTCGGAGGAGACCGGTCGCCATGCCCACT 34

||||||| || ||||| ||||| ||||| ||||| ||||| |||||

Db 278 AGGCTCGGAGGAGACCGGTCGCCATGCCCACT 311

## RESULT 11

BE758121

LOCUS BE758121 464 bp mRNA linear EST 04-OCT-2001

DEFINITION 477654 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE758121

VERSION BE758121.1

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 464)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCGACGACGACG

Plate: 102 row: L column: 5

Seq primer: ATTAGTGACACTATAG.

Location/Qualifiers

1..464

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SP08T6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from testis, thymus,

semitendonsus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

BASE COUNT 73 a 172 c 126 g 93 t

## ORIGIN

Query Match 76.5%; Score 26; DB 13; Length 464;

Best Local Similarity 85.3%; Pred. No. 28;

Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 AGGCTCCGAGGAGAACGGCTGCCATGCCAACT 34
||||||| ||||| || || ||||| ||||| |||
Db 75 AGGCTCCGAGGAGGAGCGCGCGCCATGCCGACT 108

RESULT 12
LOCUS AA234854 466 bp mRNA linear EST 06-AUG-1997
DEFINITION z177901.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:659456 5'
similar to gb:M13981 INRIBIN ALPHA CHAIN PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AA234854
VERSION AA234854.1 GI:1859384
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 466)
AUTHORS Hillier, L., Allen, M., Bowles, L., Duboue, T., Geisel, G., Jost, S.,
Kucabari, L., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
T., Waterston, R. and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 676 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..466
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:5563423"
/db_xref="taxon:9606"
/clone="IMAGE:659456"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH103"
/Note="Organ: mixed (see below); Vector: pT73D-pac
(pharmacia) with a modified polylinker; Site.1: Not f;
Site.2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NHM, pregnant uterus
NBHPU, and fetal heart NDH19W) were mixed, and ss circles
were used in vitro. Following HAP purification, this DNA
was made as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of 1.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 79 a 151 c 119 g 116 t 1 others
ORIGIN
Query Match 76.5%; Score 26; DB 9; Length 466;
Best Local Similarity 85.3%; Pred. No. 28;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGGCTCCGAGGAGAACGGCTGCCATGCCAACT 34
||||||| ||||| || ||||| ||||| |||||
Db 62 AGGCTCCGAGGAGAACGGCTGCCATGCCAACT 95

RESULT 13
LOCUS BE721322 1074 bp mRNA linear EST 05-FEB-2002
DEFINITION ACENCOUHT_6419289 NCL_CGAP_OV44 Mus musculus cDNA clone
IMAGE:5504258 5', mRNA sequence.
ACCESSION BM460913
VERSION BM460913.1 GI:18509953
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1074)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

```

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cagpts-r@mail.nih.gov  
 Tissue Procurement: Aaron Hsueh  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM12145 row: f column: 03  
 High quality sequence start: 16  
 High quality sequence stop: 739.  
 Location/Qualifiers

## FEATURES

source

1. .1074  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5504258"  
 /clone\_lib="NCI\_CGAP\_Ov44"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: ovary; PM5G-treated; Vector: pcMV-SPORT6.ccdB; Site\_1: EcoRV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Library constructed by Life Technologies. Note: this is a NCI\_CGAP Library."  
 176 a 343 c 319 g 232 t 4 others

BASE COUNT  
 ORIGIN

Query Match 76.5%; Score 26; DB 13; Length 1074;  
 Best Local Similarity 85.3%; Pred. No. 33;  
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGGCCTCCGAGAGAACCGCTGCCATGCCAACT 34  
 ||||| ||||| ||||| ||||| ||||| ||  
 DB 815 AGGCCTCCGAGAGAACCGCTGCCATGCCCTCT 848

## RESULT 15

BM461343

LOCUS

BM461343 1118 bp mRNA linear EST 05-FEB-2002  
 AGENCOURT\_6419744 NCI\_CGAP\_Ov44 Mus musculus cDNA clone  
 IMAGE:5504094 5', mRNA sequence.

ACCESSION BM461343

VERSION BM461343.1 GI:18510383

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1118)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cagpts-r@mail.nih.gov

Tissue Procurement: Aaron Hsueh

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12144 row: o column: 07

High quality sequence stop: 672.

Location/Qualifiers

source

1. .1118  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5504094"  
 /clone\_lib="NCI\_CGAP\_Ov44"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: ovary; PM5G-treated; Vector: pcMV-SPORT6.ccdB; Site\_1: EcoRV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size

2.2 kb. Library constructed by Life Technologies. Note:  
 this is a NCI\_CGAP Library."  
 BASE COUNT 179 a 372 c 323 g 244 t  
 ORIGIN

Query Match 76.5%; Score 26; DB 13; Length 1118;

Best Local Similarity 85.3%; Pred. No. 33;

Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGGCCTCCGAGAGAACCGCTGCCATGCCAACT 34

||||| ||||| ||||| ||||| ||||| ||

DB 793 AGGCCTCCGAGAGAACCGCTGCCATGCCCTCT 826

Search completed: March 11, 2003, 10:25:57

Job time : 1173.39 secs